To appear in *Quantitative Finance*, Vol. 00, No. 00, Month 20XX, 1–17

On the properties of the Lambda value at risk: robustness, elicitability and consistency

M. BURZONI*, I. PERI† and C. M. RUFFO‡

∗Department of Mathematics, ETH Zürich, Rämistrasse 101, 8092 Zürich, Switzerland
†Department of Finance, University of Greenwich, 30 Park Row, London SE10 9LS, UK
‡Department of Statistics and Quantitative Methods, University of Milano Bicocca, Via Bicocca degli Arcimboldi 8, 20126 Milan, Italy

(*v1 released April 2016*)

Recently, financial industry and regulators have enhanced the debate on the good properties of a risk measure. A fundamental issue is the evaluation of the quality of a risk estimation. On the one hand, a backtesting procedure is desirable for assessing the accuracy of such an estimation and this can be naturally achieved by elicitable risk measures. For the same objective, an alternative approach has been introduced by Davis (2016) through the so-called consistency property. On the other hand, a risk estimation should be less sensitive with respect to small changes in the available data set and exhibit qualitative robustness. A new risk measure, the Lambda value at risk (ΛVaR), has been recently proposed by Frittelli *et al.* (2014), as a generalization of VaR with the ability to discriminate the risk among P&L distributions with different tail behaviour. In this article, we show that ΛVaR also satisfies the properties of robustness, elicitability and consistency under some conditions.

Keywords: Consistency; Elicitability; Lambda Value at Risk; Law invariant risk measures; Risk measures; Robustness

JEL Classification: C13, D81, G17

1. Introduction

Risk measurement is a matter of primary concern to the financial services industry. The most widely used risk measure is the value at risk (VaR), which is the negative of the right λ-quantile q^+_λ, for some conventional confidence level λ (e.g. 1%). VaR became popular as a law invariant risk measure for its simple formulation and facility of computation, however, it presents several limits. First, VaR lacks convexity with respect to random variables which, in general, penalize diversification. VaR satisfies, instead, the quasi-convexity property with respect to distributions (Drapeau and Kupper 2012, Frittelli *et al.* 2014). This condition has a natural interpretation in terms of compound lotteries: the risk of the compound lottery is not higher than the one of the riskiest lottery. Another relevant issue of VaR is the lack of sensitivity to the tail risk as it attributes the same risk to distributions having the same quantile but different tail behaviour.

Recently, a new risk measure, the Lambda value at risk (ΛVaR), has been proposed by Frittelli *et al.* (2014). ΛVaR seems to be interesting for its ability to capture the tail risk by generalizing

*Email: matteo.burzoni@math.ethz.ch
†Corresponding author. Email: i.peri@greenwich.ac.uk
‡Email: c.ruffo@campus.unimib.it

VaR. Specifically, ΛVaR is defined as follows:

$$\Lambda VaR(F) := -\inf\{x \in \mathbb{R} : F(x) > \Lambda(x)\}$$

where $\Lambda : \mathbb{R} \to [\lambda^m, \lambda^M]$ with $0 < \lambda^m \leq \lambda^M < 1$ is a right continuous and monotone function. When the Λ function is constantly equal to some $\lambda \in (0,1)$ it coincides with the definition of VaR with confidence level λ. The main idea is that the confidence level can change and it is a function of the asset's losses. In this way, ΛVaR is able to discriminate the risk among P&L distributions with the same quantile but different tail behaviour. In this regard, the sensitivity of ΛVaR is up to the λ^m-quantile of a distribution, since, by definition, $\Lambda VaR(F) \leq VaR_{\lambda^m}(F)$. Nevertheless, the requirement $\lambda^m > 0$ is only technical and λ^m can be chosen arbitrarily close to 0. Properties of ΛVaR such as monotonicity and quasiconvexity are obtained in (Frittelli *et al.* 2014) in full generality (i.e. allowing also for $\lambda^m = 0$).

The purpose of this paper is to study if ΛVaR satisfies other important properties for a risk measure also satisfied by VaR. We first focus on the so-called *robustness* that refers to the insensitivity of a risk estimator to small changes in the data set. We adopt the Hampel's classical notion of qualitative robustness (Hampel *et al.* 1986, Huber 1981), also considered by Cont *et al.* (2010) for general risk measures (a stronger notion has been later proposed by Krätschmer *et al.* (2014) for convex risk measures). We show that the historical estimator of ΛVaR is robust within a family of distributions which depends on Λ. In particular, we recover the result of Cont *et al.* (2010) for VaR, in the case of $\Lambda \equiv \lambda \in (0,1)$.

A second property we investigate is the *elicitability* for ΛVaR . Several authors underlined the importance of this property in the risk management and backtesting practice (Gneiting 2011, Ziegel 2014, Embrechts and Hofert 2014, Bellini and Bignozzi 2015). Specifically, the elicitability allows the comparison of risk measure forecasts and provides a natural methodology to perform the backtesting. As for the case of VaR, also ΛVaR is elicitable in a particular family of distributions which depends on Λ and, for the particular case of $\Lambda \equiv \lambda \in (0,1)$, we recover the results of Gneiting (2011). Note that the elicitability for Λ decreasing was already observed by Bellini and Bignozzi (2015), we extend here to the most interesting case of Λ increasing.

Finally, we study the consistency property, as recently proposed by Davis (2016)[1]. In this study, Davis argues that the decision-theoretic framework of elicitability assumes the strong assumption that the theoretical P&L distribution is known and remain unchanged at any time. He thus suggests, under a more refined framework, the use of the so-called *consistency* property, in order to verify if a risk measure produces accurate estimates.We show that ΛVaR satisfies the consistency property without any assumption on the P&L generating process, as in the case of VaR.

The structure of the paper is as follows. After introducing the basic notions and definitions, in Section 2, we start examining the robustness property in Section 3. We dedicate the Section 4 to the elicitability of ΛVaR. Finally, in Section 5, we refine the theoretical framework and we verify the consistency of ΛVaR.

2. Notations and definitions

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a non-atomic probability space and $L^0 := L^0(\Omega, \mathcal{F}, \mathbb{P})$ be the space of $\mathcal{F}$-measurable random variables that are $\mathbb{P}$-almost surely finite. We assume that $X \in L^0$ represents a financial position (i.e. a loss when $X < 0$ and a profit when $X > 0$). Any random variable $X \in L^0$ induces a probability measure P_X on $(\mathbb{R},\mathcal{B}_{\mathbb{R}})$ by $P_X(B) = \mathbb{P}(X^{-1}(B))$ for every Borel set $B \in \mathcal{B}_{\mathbb{R}}$ and $F(x) := P_X(-\infty, x]$ denote its distribution function. Let $\mathcal{D} := \mathcal{D}(\mathbb{R})$ be the set of distribution functions and $\mathcal{D}_1$ those with finite first moment.

[1]During the review process of this paper the consistency property has been renamed by Davis as *calibration* of predictions in a dynamic setting.

A risk measure is a map $\rho : L \subseteq L^0 \to \overline{\mathbb{R}}$ that assigns to each return $X \in L$ a number representing the minimal amount of capital required by the regulator in order to cover its financial risk. The majority of risk measures used in finance are distribution-based risk measures, that is, they assign the same value to random variables with the same distribution. Such risk measures ρ are called law-invariant, more formally they satisfy:

$$X \sim_d Y \Rightarrow \rho(X) = \rho(Y).$$

In this way, a risk measure ρ can be represented as a map on a set $\mathcal{M} \subseteq \mathcal{D}$ of distributions. With a slight abuse of notation, we still denote this map by ρ and set:

$$\rho(F) := \rho(X)$$

where F is the distribution function of X. Since the seminal paper by Artzner *et al.* (1999), the theory of risk measures has been based on the study of their minimal properties. Also when risk measures are defined on distributions, monotonicity is generally accepted; formally, for any $F_1, F_2 \in \mathcal{M}$, ρ is monotone if:

$$F_1(x) \geq F_2(x),\ \forall x \in \mathbb{R} \text{ implies } \rho(F_1) \leq \rho(F_2).$$

Other properties have been discussed by academics. As pointed out in Frittelli *et al.* (2014), the convexity property, for risk measures defined on distributions, is not compatible with the translation invariance property. Thus, we might require ρ to satisfy quasiconvexity (Drapeau and Kupper 2012, Frittelli *et al.* 2014):

$$\text{for any } \gamma \in [0,1], \quad \rho(\gamma F_1 + (1-\gamma)F_2) \leq \max(\rho(F_1), \rho(F_2)).$$

It is widely accepted in the financial industry to adopt the risk measure Value at Risk (VaR) at a confidence level $\lambda \in (0,1)$, that is defined as follows (see Artzner *et al.* 1999, Definition 3.3):

$$VaR_\lambda(F) := -\inf\{x \in \mathbb{R} : F(x) > \lambda\}. \tag{1}$$

VaR is monotone and quasiconvex (Frittelli *et al.* 2014) but, obviously from the definition, it is not tail-sensitive. In order to overcome its limits, the Basel Committee (2013) recommends the use of Expected Shortfall (ES), formally given by:

$$ES_\lambda(F) := \frac{1}{\lambda}\int_0^\lambda VaR_s(F)ds. \tag{2}$$

ES is able, by definition, to evaluate the tail risk and it satisfies the subadditivity property on random variables (Artzner *et al.* 1999).

Another tail sensitive risk measure is Lambda Value at Risk (ΛVaR), recently introduced by Frittelli *et al.* (2014), whose properties are the main topic of this paper. ΛVaR generalizes VaR by considering a function Λ instead of a constant λ in the definition of VaR. The advantages of considering the Λ function are twofold: on the one hand, ΛVaR provides a criterion to change the confidence level when the market condition changes (e.g. putting aside more capital in case of expected greater losses), on the other hand, it allows differentiating the risk of P&L distributions with different tail behaviour. Formally, ΛVaR is defined by:

Definition 1

$$\Lambda VaR(F) := -\inf\{x \in \mathbb{R} : F(x) > \Lambda(x)\} \tag{3}$$

where $\Lambda : \mathbb{R} \to [\lambda^m, \lambda^M]$ with $0 < \lambda^m \leq \lambda^M < 1$ is a right continuous and monotone function.

Intuitively, if both F and Λ are continuous, ΛVaR is given by the smallest intersection between F and Λ. Unlike ES, ΛVaR lacks subadditivity, positive homogeneity and translation invariance when defined on random variables, nevertheless, ΛVaR is monotone and quasiconvex on the set of distributions (for a discussion on these properties see Section 4 of Frittelli *et al.* 2014).

3. Robustness

Evaluating the goodness of a risk measure involves determining how its computation can be affected by estimation issues. The problem consists in examining the sensitivity of a risk measure to small changes in the available data set; for this reason, *robustness* seems to be a key property. In this context, the first rigorous study is given by Cont *et al.* (2010). The authors pointed out that the notion of robustness should be referred to the "risk estimator", as outcome of a "risk measurement procedure" (see Cont *et al.* 2010, for details), and they founded the problem on the Hampel's classical notion of qualitative robustness (Hampel *et al.* 1986, Huber 1981). Basically, a risk estimator is called robust if small changes in the P&L distribution imply small changes in the law of the estimator. They consider the case of historical estimators $\hat{\rho}^h$, those obtained by applying the risk measure ρ to the empirical distribution $\hat{F}$, and they conclude that historical estimators of VaR lead to more robust procedures than alternative law-invariant coherent risk measures.

Afterwards, Krätschmer *et al.* (2012) and Krätschmer *et al.* (2014) argued that the Hampel's notion does not discriminate among P&L distributions with different tail behaviour and, hence, is not suitable for studying the robustness of risk measures that are sensitive to the tails, such as ES. So they focused on the case of law-invariant coherent risk measures and they showed that robustness is not entirely lost, but only to some degree, if a stronger notion is used.

Substantially, the robustness of a risk estimator is based on the choice of a particular metric and different metrics leads to a more or less strong definition. However, as pointed out by Embrechts *et al.* (2014), a proper definition of robustness is still a matter of primary concern. The aim of this section is to study the robustness of ΛVaR, where we use the weakest definition of robustness proposed by Cont *et al.* (2010).

Let us denote with $\mathbf{x} \in \mathcal{X}$ the n-tuple representing a particular data set, where $\mathcal{X} = \cup_{n\geq 1}\mathbb{R}^n$ is the set of all the possible data sets. The estimation of F given a particular data set $\mathbf{x}$ is denoted with $\hat{F}$ and represents the map $\hat{F} : \mathcal{X} \to \mathcal{D}$. We call risk estimator the map $\hat{\rho} : \mathcal{X} \to \mathbb{R}$ that associates to a specific data set $\mathbf{x}$ the following value:

$$\hat{\rho}(\mathbf{x}) := \rho(\hat{F}(\mathbf{x})).$$

In particular, the historical estimator $\hat{\rho}^h$ associated to a risk measure ρ is the estimator obtained by applying ρ to the empirical P&L distribution, F^{emp}, defined by $F^{emp}(x) := \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{(x\geq x_i)}$ with $n \geq 1$, that is:

$$\hat{\rho}^h(\mathbf{x}) := \rho(F^{emp}(\mathbf{x})).$$

Let us denote with $d(\cdot,\cdot)$ the Lévy metric, such that for any two distributions $F, G \in \mathcal{D}$ we have

$$d(F, G) := \inf\{\varepsilon > 0 \mid F(x - \varepsilon) - \varepsilon \leq G(x) \leq F(x + \varepsilon) + \varepsilon \quad \forall x \in \mathbb{R}\}.$$

Hereafter, we recall the definition of $\mathcal{C}$-robustness of a risk estimator as proposed by Cont *et al.* (2010), where $\mathcal{C}$ is a subset of distributions.

Definition 2 (Cont *et al.* 2010) A risk estimator $\hat{\rho}$ is $\mathcal{C}$-robust at F if for any $\varepsilon > 0$ there exists

$\delta > 0$ and $n_0 > 1$ such that, for all $G \in \mathcal{C}$:

$$d(G,F) \leq \delta \Rightarrow d(\mathcal{L}_n(\hat{\rho}, G), \mathcal{L}_n(\hat{\rho}, F)) \leq \varepsilon \quad \forall n \geq n_0$$

where d is the Lévy distance and $\mathcal{L}_n(\hat{\rho}, F)$ is the law of the estimator $\rho(\hat{F}(\mathbf{X}))$ with $\mathbf{X} := (X_1, \dots, X_n)$ a vector of independent random variables with common distribution F.

As a consequence of a generalization of the Hampel's theorem, Cont *et al.* (2010) obtained the following result:

Corollary 3 (Cont *et al.* 2010) If a risk measure ρ is continuous in $\mathcal{C}$ respect to the Lévy metric, then the historical estimator, $\hat{\rho}^h$ is $\mathcal{C}$-robust at any $F \in \mathcal{C}$.

Hence, they show that the historical estimator of VaR_λ is robust with respect to the following set:

$$\mathcal{C}_\lambda := \left\{F \in \mathcal{D} \mid q_\lambda^-(F) = q_\lambda^+(F)\right\} \tag{4}$$

where $q_\lambda^+(F) := \inf\{x \mid F(x) > \lambda\}$ and $q_\lambda^-(F) := \inf\{x \mid F(x) \geq \lambda\}$. Substantially, when the quantile of the true P&L distribution is unique, then the empirical quantile is robust. In addition, they showed that the historical estimator of ES_λ is not robust. More important, they pointed out a conflict between convexity (on random variables) and robustness: any time the convexity property is required on distribution-based risk measures, its historical estimator fails to be robust.

We use the result by Cont *et al.* (2010) in Corollary 3 to prove under which conditions the historical estimator of ΛVaR is robust.

Assumption 4 In this section we assume that $\Lambda : \mathbb{R} \mapsto [\lambda^m, \lambda^M]$ is a continuous function.

First, let us consider the following set:

$$E_F := \{x \in \mathbb{R} \mid F(x) = \Lambda(x) \text{ or } F(x^-) = \Lambda(x)\}$$

which consists of those points where the distribution F (or the left-continuous version of F) intersects Λ. We introduce the following class $\mathcal{C}_\Lambda$ of distributions:

$$\mathcal{C}_\Lambda := \{F \in \mathcal{D} \mid F((x, x+\varepsilon)) > \Lambda((x, x+\varepsilon)) \quad \text{for some } \varepsilon = \varepsilon(x) > 0,\ \forall x \in E_F\} \tag{5}$$

where $F((x, x+\varepsilon))$ and $\Lambda((x, x+\varepsilon)$ are the images of the interval $(x, x+\varepsilon)$ through F and Λ respectively. The set $\mathcal{C}_\Lambda$ consists of those distributions that do not coincide with Λ on any interval. In the special case of $\Lambda \equiv \lambda \in (0,1)$, it simply means that the quantile is uniquely determined, thus, the family $\mathcal{C}_\Lambda$ coincides with the one in (4) considered by Cont *et al.* (2010) for the robustness of VaR_λ. Note also that for Λ decreasing this condition is automatically satisfied and hence $\mathcal{C}_\Lambda = \mathcal{D}$.

In the following proposition we show that the historical estimator of ΛVaR is robust in the class $\mathcal{C}_\Lambda$ of distribution functions.

Proposition 5 ΛVaR is continuous on $\mathcal{C}_\Lambda$. Hence, $\widehat{\Lambda VaR}^h$ is $\mathcal{C}_\Lambda$-robust.

Proof. We only need to show continuity of ΛVaR respect to the Lévy metric the rest follows from Corollary 3 by Cont *et al.* (2010).
Fix $\varepsilon > 0$ and $F \in \mathcal{C}_\Lambda$. Let $\overline{x} := -\Lambda VaR(F)$. For any $n \in \mathbb{N}$, define the sets $A_n := \{x \in (-\infty, \overline{x} - \varepsilon] \mid \Lambda(x) - F(x + 1/(2n)) \geq 1/n\}$. Observe that, for $x \in A_n$, we have

$$\frac{1}{n+1} \leq \frac{1}{n} \leq \Lambda(x) - F\left(x + \frac{1}{2n}\right) \leq \Lambda(x) - F\left(x + \frac{1}{2(n+1)}\right)$$

and hence $A_n \subseteq A_{n+1}$. We first show that

$$(-\infty, \overline{x} - \varepsilon] = \bigcup_{n \in \mathbb{N}} A_n.$$

The inclusion $\supseteq$ is obvious. Fix $x \in (-\infty, \overline{x} - \varepsilon]$ and let $\gamma := \Lambda(x) - F(x)$. By definition of $\overline{x}$ and $\mathcal{C}_\Lambda$ we have that $\Lambda(x) > F(x)$ and hence $\gamma > 0$. From the right-continuity of $\Lambda - F$, and the continuity of Λ, for any $\varepsilon' > 0$ there exists $n_0 \in \mathbb{N}$ such that $\forall n \geq n_0$, $\Lambda(x + 1/(2n)) - F(x + 1/(2n)) \geq \gamma - \varepsilon'$ and $\Lambda(x) - \Lambda(x + 1/(2n)) \geq -\varepsilon'$. Take now $\varepsilon' = \gamma/4$ to obtain

$$\Lambda(x) - F(x+1/(2n)) = \Lambda(x+1/(2n)) - F(x+1/(2n)) + \Lambda(x) - \Lambda(x+1/(2n)) \geq \gamma - \gamma/4 - \gamma/4 = \gamma/2$$

since $\gamma > 0$, for a sufficiently large n we get $\Lambda(x) - F(x + 1/(2n)) \geq 1/n$ and hence $x \in A_n$ for some $n \in \mathbb{N}$, as claimed.

We now show that there exists $n_0 \in \mathbb{N}$ such that

$$(-\infty, \overline{x} - \varepsilon] = \bigcup_{n=1}^{n_0} A_n.$$

If indeed $A_{n+1} \setminus A_n \neq \emptyset$ for infinitely many $n \in \mathbb{N}$, then there exists a convergent subsequence $\{x_k\}$ with $x_k \in A_{n_k+1} \setminus A_{n_k}$ and $\tilde{x} := \lim_{k\to\infty} x_k$ such that: i) $-\infty < \tilde{x} \leq \overline{x} - \varepsilon$ and ii) $F(\tilde{x}) \geq \Lambda(\tilde{x})$. i) follows from the fact that Λ has a lower bound λ^m while F obviously tends to 0 as x approaches $-\infty$. There exists therefore $M > 0$ and n_M such that $(-\infty, M] \subseteq A_n$ for every $n \geq n_M$; ii) follows from $x_k \notin A_{n_k}$ which implies $\Lambda(x_k) - F(x_k + 1/(2n_k)) < 1/n_k$ and the right-continuity of F which implies

$$F(\tilde{x}) \geq \limsup F(x_k + 1/(2n_k)) \geq \limsup \Lambda(x_k) - 1/n_k = \Lambda(\tilde{x})$$

where the last inequality follows from the continuity of Λ. If $F(\tilde{x}) = \Lambda(\tilde{x})$, by definition of $\mathcal{C}_\Lambda$ we obtain $-\Lambda VaR(F) \leq \tilde{x} \leq \overline{x} - \varepsilon$ which is a contradiction. The same conclusion obviously follows when $F(\tilde{x}) > \Lambda(\tilde{x})$.

We have therefore shown the existence of $n_0 \in \mathbb{N}$ such that $\Lambda(x) - F(x + 1/(2n_0)) \geq 1/n_0$ for every $x \in (-\infty, \overline{x} - \varepsilon]$ Take now $\delta_1 := 1/(2n_0)$ and $G \in \mathcal{C}_\Lambda$ such that $d(F, G) < \delta_1$. We thus have, for any $x \leq \overline{x} - \varepsilon$,

$$\Lambda(x) - G(x) \geq \Lambda(x) - F(x + \delta_1) - \delta_1 \geq \frac{1}{n_0} - \delta_1 = \frac{1}{2n_0} > 0.$$

It follows

$$\Lambda VaR(G) \leq \Lambda VaR(F) + \varepsilon \tag{6}$$

which is the upper semi-continuity.
By showing the lower semi-continuity we conclude the proof. From Definition 1, for any $\varepsilon > 0$, there exists $\hat{x} \in [\overline{x}, \overline{x} + \varepsilon]$ such that $\gamma := F(\hat{x}) - \Lambda(\hat{x}) > 0$. Since Λ is continuous, there exists $\delta > 0$ such that for all $\delta' \leq \delta$, $\Lambda(\hat{x}) - \Lambda(\hat{x} + \delta') \geq -\gamma/4$. Take now $\delta_2 \leq \min\{\delta, \gamma/4, \varepsilon\}$ so that $\hat{x} + \delta_2 \in [\overline{x}, \overline{x} + \varepsilon]$. By observing that, for $G \in \mathcal{C}_\Lambda$ with $d(F, G) < \delta_2$ we have

$$G(\hat{x} + \delta_2) - \Lambda(\hat{x} + \delta_2) \geq F(\hat{x}) - \delta_2 - \Lambda(\hat{x} + \delta_2) \geq F(\hat{x}) - \Lambda(\hat{x}) - \gamma/4 - \delta_2 \geq \gamma/2$$

we obtain

$$\Lambda VaR(G) \geq -\hat{x} - \delta_2 \geq \Lambda VaR(F) - \varepsilon. \tag{7}$$

By taking $\delta := \min\{\delta_1, \delta_2\}$ and combining (6) and (7), we have that

$$\forall G \in \mathcal{C}_\Lambda \quad \text{with } d(F,G) < \delta \Longrightarrow |\Lambda VaR(F) - \Lambda VaR(G)| < \varepsilon$$

as desired. □

The Λ function adds flexibility to ΛVaR, however, when robustness is required, ΛVaR should be constructed as suggested by the set $\mathcal{C}_\Lambda$. The Λ function has to be chosen continuous and, on any interval, it cannot coincide with any distribution F under consideration. We refer to Example 11 to show how this condition can be guaranteed given a set of normal distributions of P&Ls.

4. Elicitability

The importance of this property from a financial risk management perspective has been highlighted by Embrechts and Hofert (2014) as a consequence of the surprising results obtained by Gneiting (2011) and Ziegel (2014). Indeed, Embrechts and Hofert (2014) pointed out that the elicitability allows the assessment and the comparison of risk measure forecasting estimations and a straightforward backtesting.

The term *elicitable* has been introduced by Lambert *et al.* (2008) but the general notion dates back to the pioneering work of Osband (1985). In accordance with some parts of the literature, we introduce the notation $T : \mathcal{M} \subseteq \mathcal{D} \to 2^{\mathbb{R}}$ to describe a set-valued statistical functional. Let us denote with $S(x, y)$ the realized forecasting error between the ex-ante prediction $x \in \mathbb{R}$ and the ex-post observation $y \in \mathbb{R}$, where S is a function $S : \mathbb{R} \times \mathbb{R} \to [0, +\infty)$ called "scoring" or "loss". According to Gneiting (2011) a scoring function S is consistent for the functional T if

$$\mathbb{E}_F[S(t, Y)] \leq \mathbb{E}_F[S(x, Y)] \tag{8}$$

for all F in $\mathcal{M}$, all $t \in T(F)$ and all $x \in \mathbb{R}$. It is strictly consistent if it is consistent and equality of the expectations implies that $x \in T(F)$.

Definition 6 (Gneiting 2011) A set-valued statistical functional $T : \mathcal{M} \to 2^{\mathbb{R}}$ is elicitable if there exists a scoring function S that is strictly consistent for it.

Bellini and Bignozzi (2015) have recently proposed a slightly different definition of elicitability. They consider only single-valued statistical functionals as a natural requirement in financial applications. In addition, they adopt additional properties for the scoring function. We also consider single-valued statistical functionals but without imposing any restriction on the scoring function.

Definition 7 A statistical functional $T : \mathcal{M} \to \mathbb{R}$ is elicitable if there exists a scoring function S such that

$$T(F) = \arg\min_x E_F[S(x, Y)] \quad \forall F \in \mathcal{M}. \tag{9}$$

Definition 6 restricted to the case of single-valued statistical functional is equivalent to Definition 7 when the minimum is unique. The statistical functional associated to a risk measure is the map $T : \mathcal{M} \to \overline{\mathbb{R}}$ such that $T(F) = -\rho(F)$ for any distribution F. We adopt this sign convention in accordance with part of the literature. We say that a risk measure is elicitable if the associated statistical functional T is elicitable. In the following we will restrict to $\mathcal{M} \subseteq \mathcal{D}_1$ in order to have a finite expectation of the considered scoring functions.

The statistical functional associated to VaR, $T(F) := q^+_\lambda(F)$, is elicitable on the following set:

$$\mathcal{M}_\lambda := \{F \in \mathcal{D}_1 : F \text{ strictly increasing } \} \subseteq \mathcal{C}_\lambda$$

with $\mathcal{C}_\lambda$ as in (4), and with the following scoring function (Gneiting 2011):

$$S(x, y) = \lambda(y - x)^+ + (1 - \lambda)(y - x)^-. \tag{10}$$

Let us denote with $T_\Lambda : \mathcal{D} \to \mathbb{R}$ the statistical functional associated to ΛVaR such that:

$$T_\Lambda(F) = -\Lambda VaR(F) \tag{11}$$

and consider the set $\mathcal{M}_\Lambda \subseteq \mathcal{D}_1$ defined as follows:

$$\mathcal{M}_\Lambda = \{F \in \mathcal{D}_1 : \exists\, \bar{x} \text{ s.t. } \forall x < \bar{x},\ F(x) < \Lambda(x) \text{ and } \forall x > \bar{x},\ F(x) > \Lambda(x)\}. \tag{12}$$

Once again this set coincides with $\mathcal{M}_\lambda$ when $\Lambda \equiv \lambda$. In Bellini and Bignozzi (2015) it has been shown that ΛVaR is elicitable under a stronger definition of elicitability and for the special case of Λ continuous and decreasing. In the next theorem we prove that ΛVaR is elicitable using the general Definition 7 and under less restrictive conditions on Λ. Specifically, we show that ΛVaR is elicitable on the particular class of distribution $\mathcal{M}_\Lambda$ in (12) depending on Λ.

Theorem 8 For any monotone and right continuous function $\Lambda : \mathbb{R} \to [\lambda^m, \lambda^M]$, with $0 < \lambda^m \leq \lambda^M < 1$, the statistical functional $T_\Lambda : \mathcal{D} \to \mathbb{R}$ defined in (11) is elicitable on the set $\mathcal{M}_\Lambda \subseteq \mathcal{D}_1$ defined in (12) with a loss function given by

$$S(x, y) = (y - x)^- - \int_y^x \Lambda(t)dt. \tag{13}$$

Proof. We need to prove that

$$T(F) = \arg\min_x \int_{\mathbb{R}} S(x, y)dF(y).$$

In order to find a global minimum we first calculate the left and right derivatives of $\int_{\mathbb{R}} S(x, y)dF(y)$. Applying dominated convergence theorem we obtain:

$$\begin{aligned}
\frac{\partial^-}{\partial x}\int_{\mathbb{R}} S(x, y)dF(y) &= \frac{\partial^-}{\partial x}\int_{\mathbb{R}} \Big((y - x)^- - \int_y^x \Lambda(t)dt\Big)dF(y) \\
&= \int_{\mathbb{R}} \Big(\frac{\partial^-}{\partial x}(y - x)^- - \frac{\partial^-}{\partial x}\int_y^x \Lambda(t)dt\Big)dF(y) \\
&= \int_{\mathbb{R}} \Big(\mathbf{1}_{(y<x)} - \Lambda(x^-)\Big)dF(y) \\
&= \lim_{t \uparrow x} F(t) - \Lambda(x^-) = F(x^-) - \Lambda(x^-).
\end{aligned}$$

Analogously for the right derivative

$$\begin{aligned}\frac{\partial^+}{\partial x}\int_{\mathbb{R}} S(x,y)dF(y) &= \int_{\mathbb{R}} \Big(\mathbf{1}_{(y\le x)} - \Lambda(x)\Big)dF(y)\\ &= F(x) - \Lambda(x).\end{aligned}$$

Observe now that $x^* = \inf\{x \in \mathbb{R} : F(x) > \Lambda(x)\}$, that is the statistical functional associated to ΛVaR, satisfies, for every $F \in \mathcal{M}_\Lambda$,

$$\begin{aligned}&\forall x < x^* \quad F(x) < \Lambda(x), \quad F(x^-) \le \Lambda(x^-)\ ;\\ &\forall x > x^* \quad F(x) > \Lambda(x), \quad F(x^-) \ge \Lambda(x^-)\ ;\end{aligned} \tag{14}$$

from which we deduce

$$\begin{aligned}&\forall x < x^* \quad \frac{\partial^-}{\partial x}\int_{\mathbb{R}} S(x,y)dF(y) \le 0, \quad \frac{\partial^+}{\partial x}\int_{\mathbb{R}} S(x,y)dF(y) < 0;\\ &\forall x > x^* \quad \frac{\partial^-}{\partial x}\int_{\mathbb{R}} S(x,y)dF(y) \ge 0, \quad \frac{\partial^+}{\partial x}\int_{\mathbb{R}} S(x,y)dF(y) > 0.\end{aligned} \tag{15}$$

This implies that x^* is a local minimum. By showing that there are no other local minima we obtain that x^* is the unique global minimum. Take first $x < x^*$. Observe that, by applying dominated convergence theorem, $I(x) := \int_{\mathbb{R}} S(x,y)dF(y)$ is a continuous function. Moreover, I is not constant on any interval in $(-\infty, x^*]$ since, from (15), we have $\frac{\partial^+}{\partial x} I < 0$. Since I is continuous and, from (15), the left and right derivatives are non-positive, we have that any sequence converging to x^- is decreasing. Analogously, any sequence converging to x^+ is increasing. In other words, there exists $\delta > 0$ such that, $I(x_1) > I(x) > I(x_2)$ for all $x - \delta < x_1 < x < x_2 < x + \delta$. Thus x is not a local minimum. The case $x > x^*$ is analogous. We can conclude that ΛVaR is elicitable on the class of probability measures $\mathcal{M}_\Lambda$ defined in (12). □

Remark 9 It is easy to prove that the scoring function in (13) can be rewritten as follows:

$$S(x,y) = \frac{\int_y^x \Lambda(t)dt}{x-y}(y-x)^+ + \left(1 - \frac{\int_y^x \Lambda(t)dt}{x-y}\right)(y-x)^-. \tag{16}$$

if $x \ne y$ and $S(x,x) = 0$. It is evident the similarity with the scoring function of VaR in (10). Moreover, note that If Λ is non-increasing obviously (12) is satisfied by every $F \in \mathcal{D}_1$ increasing so that we recover the result of Bellini and Bignozzi (2015).

In general, the elicitability of ΛVaR using the scoring function (13) requires that Λ is crossed only once by any possible F at the level $\bar{x} = -\Lambda VaR(F)$ as shown in (12).

Remark 10 ΛVaR with a decreasing function Λ is elicitable on the set of all the distributions. In this case, $\mathcal{M}_\Lambda \equiv \mathcal{D}_1$, since F is non-decreasing and the derivatives of Λ are negative. When Λ is non-increasing, ΛVaR is elicitable on the set of increasing distribution functions.

If we additionally require continuity of Λ we observe that $\mathcal{M}_\Lambda \subseteq \mathcal{C}_\Lambda$ where $\mathcal{C}_\Lambda$ is defined in (5). This implies that the set of distributions where ΛVaR is elicitable guarantees also that ΛVaR is robust. Hereafter, we provide an example of a construction of ΛVaR with non-decreasing Λ that is elicitable and robust given a set of normal distributions of P&Ls.

Example 11 Denote by $\Phi(x)$ the distribution function of a standard normal distribution. Let $\mathcal{M} := \{\Phi(\frac{x-\mu_i}{\sigma_i})\}_{i\in I}$ for some collection I such that $\overline{\mu} := \sup \mu_i < \infty$ and $\underline{\sigma} := \inf \sigma_i > 0$. Set

$\mu > \overline{\mu}$, $0 < \sigma < \underline{\sigma}$ and define

$$\Lambda(x) := \begin{cases} \lambda^m & x \leq x^m \\ \Phi\left(\dfrac{x-\mu}{\sigma}\right) & x^m \leq x < x^M \\ \lambda^M & x \geq x^M. \end{cases}$$

If $x^m \leq x^M$ are such that $0 < \lambda_m \leq \Phi(\frac{x^m-\mu}{\sigma})$ and $\Phi(\frac{x^M-\mu}{\sigma}) \leq \lambda^M$ then Λ is non-decreasing and continuous. Moreover, from Theorem 8, ΛVaR is elicitable on $\mathcal{M}$.

In order to have an elicitable ΛVaR with the scoring function (13) we need to build the Λ function under a certain condition that depends on the set of the P&L distributions. In particular, the scoring function (13) guarantees the elicitability of ΛVaR with non-decreasing Λ only in the class of probability measures $\mathcal{M}_\Lambda$ in (12) as shown by the following counterexample.

Example 12 Let $0 < \varepsilon < 0.5\%$. Let $\Lambda(x)$ and $F(x)$ as follows

$$F(x) = \begin{cases} 0 & x < -100 \\ 1.5\% & -100 \leq x < 4 \\ 1 & x \geq 4 \end{cases} \qquad \Lambda(x) = \varepsilon + \begin{cases} 0 & x < -101 \\ (x+101)/100 & -101 \leq x < -99 \\ 2\% & x \geq -99. \end{cases}$$

$F(x)$ is the cumulative distribution function of a random variable Y with distribution: $Y = -100$ with probability $p = 1.5\%$ and $Y = 4$ with probability $1 - p = 98.5\%$.

It is easy to compute that the statistical functional associated to ΛVaR is $T_\Lambda(F) = -100$. If ΛVaR is elicitable T_Λ should be the minimizer of

$$g(x) := \mathbb{E}[S(x,Y)] = S(x,-100)\frac{1.5}{100} + S(x,4)\frac{98.5}{100}.$$

Since S for ΛVaR is defined as in (13), we need compute the primitive for Λ that is given by

$$\Psi(t) = \int \Lambda(t) = \varepsilon t + \begin{cases} 0 & t < -101 \\ \dfrac{(t^2/2 + 101t)}{100} & -101 \leq t < -99 \\ \dfrac{2}{100}t & t \geq -99. \end{cases}$$

Hence, $\Psi(-100) = -51 - 100\varepsilon$ and $\Psi(4) = 8/100 + 4\varepsilon$, thus, we have $S(x,-100) = (-100-x)^- - \Psi(x) - 51 - 100\varepsilon$ and $S(x,4) = (4-x)^- - \Psi(x) + 8/100 + 4\varepsilon$ and

$$g(x) = -\Psi(x) + (-100-x)^- \frac{1.5}{100} + (4-x)^- \frac{98.5}{100} + c$$

where $c = (-51 - 100\varepsilon) \cdot 1.5\% + (0.08 + 4\varepsilon) \cdot 98.5\%$. Observe now that ΛVaR is not the global minimum, since $g(-100) > g(4)$. Indeed:

$$g(-100) - g(4) = -\Psi(-100) + \Psi(4) - 104 \cdot \frac{1.5}{100} = 51 + \frac{8}{100} - 104 \cdot \frac{1.5}{100} > 0.$$

We have shown that the scoring function in (13) guarantees elicitability of ΛVaR only on the set of distributions $\mathcal{M}_\Lambda$ in (12). Whether there exists another scoring function that guarantees the elicitability of ΛVaR on a larger class of distributions is an interesting question which might be object of further studies. We conclude this Section by discussing some insights on this problem and the difficulties that might arise for such an extension. In particular we investigate a necessary condition for elicitability, namely, the convex level sets property (Osband 1985).

Definition 13 If $\mathcal{M} \subseteq \mathcal{D}$ is convex we say that T has convex level sets if, for any $\gamma \in \mathbb{R}$, the level sets

$$\{T = \gamma\} := \{F \in \mathcal{M} : T(F) = \gamma\}$$

are convex, i.e. for any $\alpha \in [0,1]$ and $F_1, F_2 \in \mathcal{M}$

$$T(F_1) = T(F_2) = \gamma \Rightarrow T(\alpha F_1 + (1-\alpha)F_2) = \gamma.$$

Proposition 14 (Osband 1985) If a statistical functional $T : \mathcal{M} \subseteq \mathcal{D} \to \mathbb{R}$ is elicitable, then T has convex level sets.

Gneiting (2011) showed that ES does not satisfy this necessary condition, as a consequence, ES is not elicitable.

We have shown in Theorem 8 that ΛVaR is elicitable in $\mathcal{M}_\Lambda$, hence, it also has convex level sets in this class of distributions. The following example shows that, in general, ΛVaR might not satisfy this condition on a larger set of distributions and, thus, neither elicitability.

Example 15 Fix $0 < \varepsilon < \frac{1}{2}$ and $\lambda^M < 1$. Consider

$$F_1(x) := \sum_{k=1}^{\infty} \frac{1}{2^k} \mathbf{1}_{\left[\frac{1}{k+1}, \frac{1}{k}\right)}(x) + \varepsilon \mathbf{1}_{[0,1)} + \mathbf{1}_{[1,\infty)}$$

and

$$F_2(x) := F_1(x) + \sum_{k=1}^{\infty} (-1)^k \frac{1}{10^k} \mathbf{1}_{\left[\frac{1}{k+1}, \frac{1}{k}\right)}(x).$$

As a function Λ take $\Lambda := \varepsilon \mathbf{1}_{(-\infty,0)} + \frac{1}{2}(F_1 + F_2)\mathbf{1}_{[0,1)} + \lambda^M \mathbf{1}_{[1,\infty)}$. Observe that $\forall k \in \mathbb{N}$ $F_1(\frac{1}{2k}) > \Lambda(\frac{1}{2k})$ and $F_2(\frac{1}{2k+1}) > \Lambda(\frac{1}{2k+1})$. Moreover, $0 = F_1(x) = F_2(x) < \Lambda(x)$ for all $x < 0$. This implies $\Lambda VaR(F_1) = \Lambda Var(F_2) = 0$. Nevertheless, since $\Lambda(x) = \lambda^M < 1$ for $x \geq 1$, we have $\Lambda Var(\frac{1}{2}F_1 + \frac{1}{2}F_2) = -1$, from which the convex level set property fails.

A positive answer for the convex level sets property is given by the choice of a particular class of Λ for which the condition is satisfied on the set of increasing distribution functions.

Lemma 16 If Λ is non-decreasing and piecewise constant with a finite number of jumps, then ΛVaR has convex level sets on the set of increasing distribution functions.

Proof. We first observe that, in general, $T_\Lambda(F_i) = \gamma$ for $i = 1, 2$ implies $T_\Lambda(\alpha F_1 + (1-\alpha)F_2) \geq \gamma$ for every $\alpha \in [0,1]$ and $F_1, F_2 \in \mathcal{D}$. To this end, we prove that $\inf\{x : \alpha F_1(x) + (1-\alpha)F_2(x) > \Lambda(x)\} \geq \gamma$, with $\gamma = T_\Lambda(F_i) := \inf\{x : F_i(x) > \Lambda(x)\}$ for $i = 1, 2$. Note that by definition of $T_\Lambda(F_i)$ for $i = 1, 2$, we have $F_i(x) \leq \Lambda(x)$ for every $x \leq \gamma$. We thus get, for an arbitrary $0 \leq \alpha \leq 1$, $\alpha F_1(x) + (1-\alpha)F_2(x) \leq \Lambda(x)$ for every $x \leq \gamma$ from which $T_\Lambda(\alpha F_1 + (1-\alpha)F_2) \geq \gamma$.

For the converse inequality observe that there exists $\varepsilon > 0$ such that Λ is constant on $[\gamma, \gamma + \varepsilon)$. Since $\gamma = \inf\{x : F_i(x) > \Lambda(x)\}$ and F_i is non-decreasing, for $i = 1, 2$, then $\alpha F_1(x) + (1-\alpha)F_2(x) > \Lambda(x)$ for every $x \in (\gamma, \gamma + \varepsilon)$ from which $T_\Lambda(\alpha F_1 + (1-\alpha)F_2) \leq \gamma$. □

In conclusion, we have observed that extending the class of distributions for which the convex level sets property holds depends heavily on the specific choice of Λ and hence it seems to necessitate a case-by-case study.

5. Consistency

In this section we refer to the notion of *consistency* recently studied by Davis (2016). Davis recognized the importance of the elicitability property in the backtesting context of risk measures, but he argued that the problem can be better addressed from a different perspective. The motivation of Davis' study relates to the difficulties of predicting the "true" distribution F of portfolio financial returns. Suppose indeed you are given the information up to time $k-1$, at time k only one realization occurs and so there is not enough information to claim if the prediction of F was correct or not. Thus, Davis introduces the notion of consistency of a risk estimator that is based on the daily comparison between the realization of the risk estimator and the realized outcome, but without consideration how the predictions were arrived at. Hence, the fundamental difference with the elicitability property is that the assumption on the model generating the conditional distribution of the portfolio returns can change at any time and one should just check if the prediction is performing well or not (see Davis 2016, for a more detailed discussion).

In this section we adopt the framework of Davis. Namely, we fix $(\Omega, \mathcal{F}, \{\mathcal{F}_k\}_{k\in\mathbb{N}})$ where $\Omega = \prod_{k=1}^{\infty} \mathbb{R}_{(k)}$ is the canonical space for a real-valued data process $Y = \{Y_k\}_{k\in\mathbb{N}}$; $\mathcal{F}$ is the product sigma-algebra generated by the Borel sigma-algebra in each copy of $\mathbb{R}$ (denoted by $\mathbb{R}_{(k)}$); $\{\mathcal{F}_k\}_{k\in\mathbb{N}}$ is the natural filtration of the process Y and $\mathcal{F}_0$ the trivial sigma-algebra. The class of possible models, for this data process, is represented by a collection $\mathcal{P}$ of probability measures denoted by $\mathcal{P} := \{\mathbb{P}^\alpha, \alpha \in \mathfrak{A}\}$, where $\mathfrak{A}$ is an arbitrary index set. We denote with $\mathbb{E}^\alpha$ the expectation with respect to $\mathbb{P}^\alpha$. For every $\mathbb{P}^\alpha$ it is possible to define, for each $k \geq 1$, the conditional distribution of the random variable Y_k given $\mathcal{F}_{k-1}$, as a map $F_k^\alpha : \mathbb{R} \times \Omega \mapsto [0,1]$ satisfying: for $\mathbb{P}^\alpha$-a.e. ω, $F_k^\alpha(\cdot, \omega)$ is a distribution function, and for every $x \in \mathbb{R}$, $F_k^\alpha(x) = \mathbb{P}^\alpha(Y_k \leq x|\mathcal{F}_{k-1})$ $\mathbb{P}^\alpha$-a.s.

Definition 17 (Davis 2016) Let $\mathfrak{B}(\mathcal{P})$ be a set of strictly increasing predictable processes $b = \{b_n\}_{n\in\mathbb{N}}$ such that $\lim_{n\to\infty} b_n = \infty$ $\mathbb{P}^\alpha$-a.s. for every $\alpha \in \mathfrak{A}$, and $l : \mathbb{R}^2 \to \mathbb{R}$ a calibration function, that is a measurable function such that $\mathbb{E}^\alpha[l(T(F_k^\alpha), Y_k)|\mathcal{F}_{k-1}] = 0$ for all $\mathbb{P}^\alpha \in \mathcal{P}$. A risk measure ρ is $(l, b, \mathcal{P})$-consistent if the associated statistical functional T satisfies

$$\lim_{n\to\infty} \frac{1}{b_n} \sum_{k=1}^{n} l(T(F_k^\alpha), Y_k) = 0 \quad \mathbb{P}^\alpha\text{-a.s. } \forall \mathbb{P}^\alpha \in \mathcal{P}. \tag{17}$$

Denote by $\mathfrak{P}$ the set of all probability measures and define:

$$\mathcal{P}^0 = \{\mathbb{P}^\alpha \in \mathfrak{P} : \forall k \ F_k^\alpha(x, \omega) \ \text{ is continuous in } x \text{ for } \mathbb{P}^\alpha\text{-almost all } \omega \in \Omega\}.$$

Davis (2016) showed that VaR satisfies this consistency property for a large class of processes $\mathcal{B}(\mathcal{P})$ and for the large class of data models $\mathcal{P}^0$ with the following calibration function:

$$l(x, y) = \lambda - \mathbf{1}_{(y\leq x)}.$$

The statistical functional associated to ΛVaR is given by (11), hence we define for every k and $\alpha \in \mathfrak{A}$:

$$T_\Lambda(F_k^\alpha) := \inf\{x \mid F_k^\alpha(x) > \Lambda(x)\}.$$

Notice that $\{T_\Lambda(F_k^\alpha)\}_{k\in\mathbb{N}}$ and $\{\Lambda(T_\Lambda(F_k^\alpha))\}_{k\in\mathbb{N}}$ are predictable process, as shown in the following lemma.

Lemma 18 For every $k \geq 1$, $T_\Lambda(F_k^\alpha)$ and $\Lambda(T_\Lambda(F_k^\alpha))$ are $\mathcal{F}_{k-1}$-measurable random variables.

Proof. Fix a probability $\mathbb{P}^\alpha$ with $\alpha \in \mathfrak{A}$. Notice first that for any $y \in \mathbb{R}$, for $\mathbb{P}^\alpha$ a.e. ω, we have

$$\begin{aligned} T_\Lambda(F_k^\alpha) \geq y &\iff F_k^\alpha(x) \leq \Lambda(x) \qquad \forall x \leq y \\ &\iff F_k^\alpha(q) \leq \Lambda(q) \ \forall q \in \mathbb{Q},\ q \leq y \end{aligned}$$

where the last equivalence follows from the right-continuity of F_k^α and Λ. We therefore have

$$\{\omega \mid T_\Lambda(F_k^\alpha) \geq y\} = \bigcap_{q \in \mathbb{Q} \cap (-\infty, y]} \{\omega \mid F_k^\alpha(q) \leq \Lambda(q)\} \in \mathcal{F}_{k-1}$$

from which $T_\Lambda(F_k^\alpha)$ is an $\mathcal{F}_{k-1}$-measurable random variable.

$\Lambda(T_\Lambda(F_k^\alpha))$ is also $\mathcal{F}_{k-1}$-measurable: since Λ is right-continuous $\Lambda(x) \geq y$ iff $x \geq \Lambda^-(y)$ where $\Lambda^-(y) := \inf\{x \in \mathbb{R} \mid \Lambda(x) \geq y\}$ is the generalized inverse (Embrechts and Hofert 2013, Proposition 1) and thus

$$\{\omega \mid \Lambda(T_\Lambda(F_k^\alpha)) \geq y\} = \{\omega \mid T_\Lambda(F_k^\alpha) \geq \Lambda^-(y)\} \in \mathcal{F}_{k-1}.$$

□

By following the methodology suggested by Davis (2016), we are able to show that ΛVaR is consistent for the large class of data models $\mathcal{P}^0$, as shown in the following theorem.

Theorem 19 For each $\mathbb{P}^\alpha \in \mathcal{P}^0$,

$$\frac{1}{n}\sum_{k=1}^{n} \Lambda(T_\Lambda(F_k^\alpha)) - \mathbf{1}_{(Y_k \leq T_\Lambda(F_k^\alpha))} \to 0 \quad \mathbb{P}^\alpha\text{-a.s.} \tag{18}$$

Thus, ΛVaR is $(l, n, \mathcal{P}^0)$-consistent with

$$l(x, y) = \Lambda(x) - \mathbf{1}_{(y \leq x)}. \tag{19}$$

Before giving the proof of the theorem we show the following lemma.

Lemma 20 For each $\mathbb{P}^\alpha \in \mathcal{P}^0$,

$$\mathbb{E}^\alpha\left[\mathbf{1}_{(Y_k \leq T_\Lambda(F_k^\alpha))} \mid \mathcal{F}_{k-1}\right] = \Lambda(T_\Lambda(F_k^\alpha)), \qquad \mathbb{P}^\alpha\text{-a.s.}$$

Proof. Fix $\mathbb{P}^\alpha \in \mathcal{P}^0$. Since there is no confusion, for ease of notation, we omit the dependence on α. Observe that $U_k := F_k(Y_k)$ is uniformly distributed and

$$Y_k \leq T_\Lambda(F_k) \iff U_k \leq F_k(T_\Lambda(F_k)) = \Lambda(T_\Lambda(F_k)).$$

Note now that U_k is independent of $\mathcal{F}_{k-1}$ since, from the continuity of F_k, $\mathbb{P}(U_k \leq u_k \mid \mathcal{F}_{k-1}) = \mathbb{P}(Y_k \leq F_k^-(u_k) \mid \mathcal{F}_{k-1}) = F_k(F_k^-(u_k)) = u_k = \mathbb{P}(U_k \leq u_k)$ (where F_k^- denotes the generalized inverse of F_k). Since $\Lambda(T_\Lambda(F_k))$ is $\mathcal{F}_{k-1}$-measurable from Lemma 18, we can compute the desired conditional expectation through the application of the freezing lemma (Williams 1991, Section 9.10). Namely, define $h(x, y) := \mathbf{1}_{\{y \leq x\}}$ and let $\hat{h}(x) := \mathbb{E}[\mathbf{1}_{\{U_k \leq x\}}] = x$. Since h is a bounded

Borel-measurable function and U_k is independent of $\mathcal{F}_{k-1}$, then

$$\mathbb{E}\left[\mathbf{1}_{(Y_k \le T_\Lambda(F_k))} \mid \mathcal{F}_{k-1}\right] = \hat{h}(\Lambda(T_\Lambda(F_k))) = \Lambda(T_\Lambda(F_k))$$

where equalities are intended in the $\mathbb{P}$-a.s. sense. □

Proof of Theorem 19. Define:

$$Z_k := \Lambda(T_\Lambda(F_k^\alpha)) - \mathbf{1}_{(Y_k \le T_\Lambda(F_k^\alpha))},$$

$$S_n := \sum_{k=1}^{n} Z_k, \qquad Q_n := \sum_{k=1}^{n} (Z_k)^2, \qquad \langle S \rangle_n := \sum_{k=1}^{n} \mathbb{E}^\alpha[(Z_k)^2 \mid \mathcal{F}_{k-1}].$$

From Lemma 20, S_n is a martingale since $\mathbb{E}^\alpha[S_n - S_{n-1} \mid \mathcal{F}_{n-1}] = \mathbb{E}^\alpha[Z_n \mid \mathcal{F}_{n-1}] = 0$. We now compute $(Z_k)^2$, we use the shorthand $W := \Lambda(T_\Lambda(F_k^\alpha))$.

$$\begin{aligned}(Z_k)^2 &= \mathbf{1}_{(Y_k \le T_\Lambda(F_k^\alpha))} + W^2 - 2W\mathbf{1}_{(Y_k \le T_\Lambda(F_k^\alpha))} \\ &= W^2 \mathbf{1}_{(Y_k > T_\Lambda(F_k^\alpha))} + (1-W)^2 \mathbf{1}_{(Y_k \le T_\Lambda(F_k^\alpha))}.\end{aligned}$$

Note that since $\lambda^m \le W \le \lambda^M$ we obtain $\mathbb{E}^\alpha[(Z_k)^2] \le \max\{(\lambda^M)^2, (1-\lambda^m)^2\} < \infty$ so that S_n is a square integrable martingale. Moreover, observe that,

$$(Z_k)^2 \ge \min\{(\lambda^m)^2, (1-\lambda^M)^2\}. \tag{20}$$

Since W is $\mathcal{F}_{k-1}$-measurable, using Lemma 20,

$$\begin{aligned}\mathbb{E}^\alpha[(Z_k)^2 \mid \mathcal{F}_{k-1}] &= W^2 \mathbb{E}^\alpha[\mathbf{1}_{(Y_k > T_\Lambda(F_k^\alpha))} \mid \mathcal{F}_{k-1}] + (1-W)^2 \mathbb{E}^\alpha[\mathbf{1}_{(Y_k \le T_\Lambda(F_k^\alpha))} \mid \mathcal{F}_{k-1}] \\ &= W^2(1-W) + (1-W)^2 W \\ &= W(1-W).\end{aligned}$$

It follows that

$$\lambda^m(1-\lambda^M) \le \mathbb{E}^\alpha[(Z_k)^2 \mid \mathcal{F}_{k-1}] \le \lambda^M(1-\lambda^m)$$

which firstly implies $\langle S \rangle_n \ge n\lambda^m(1-\lambda^M) \to \infty$, and, secondly, combined with (20),

$$\frac{Q_n}{\langle S \rangle_n} \ge \frac{n \min\{(\lambda^m)^2, (1-\lambda^M)^2\}}{n\lambda^M(1-\lambda^m)} = \frac{\min\{(\lambda^m)^2, (1-\lambda^M)^2\}}{\lambda^M(1-\lambda^m)} =: \varepsilon_\alpha > 0.$$

Notice that Z_k is bounded from above by 1 for every $k \in \mathbb{N}$, thus, we have $Q_n \le n$. We can therefore conclude

$$\left|\frac{S_n}{n}\right| \le \left|\frac{S_n}{Q_n}\right| = \frac{\langle S \rangle_n}{Q_n} \left|\frac{S_n}{\langle S \rangle_n}\right| \le \frac{1}{\varepsilon_\alpha} \left|\frac{S_n}{\langle S \rangle_n}\right| \to 0 \quad \mathbb{P}^\alpha\text{-a.s.}$$

where the last term converges to 0 from Proposition 6.3 in Davis (2016). □

As a consequence of the theorem, similarly to what observed by Davis (2016) for VaR, a risk manager could use the following relative frequency measure

$$\frac{1}{n}\sum_{k=1}^{n}\Lambda(T_\Lambda(F_k^\alpha)) - \mathbf{1}_{(Y_k\leq T_\Lambda(F_k^\alpha))} \tag{21}$$

as test statistic in a finite-sample hypothesis test (as considered in Corbetta and Peri 2016). Obviously ΛVaR is also $(l, b', \mathcal{P}^0)$-consistent with $b'_n = nb_n$ and $b = \{b_n\}_n \in \mathfrak{B}(\mathcal{P})$.

Therefore, the consistency of ΛVaR, as the quantile forecasting, can be obtained under essentially no conditions on the mechanism generating the data. This is not the case of the estimates of the statistical functional T_m associated to the conditional mean (such as ES) and defined as follows:

$$T_m(F_k^\alpha) := \int_{\mathbb{R}} xF_k^\alpha(dx).$$

Indeed, Davis (2016) showed that $T_m(F_k^\alpha)$ satisfies the condition (17) with $l(x, y) = x - y$, $Q_n = \sum_{k=1}^n Z_k^2$, where $Z_k := Y_k - T_m(F_k^\alpha)$, and, remarkably, $\mathcal{P}^1 \in \mathfrak{P}$ is the set of probability measures such that:

i) for any k, $Y_k \in L^2(\mathbb{P}^\alpha)$,
ii) $\lim_{n\to\infty}\langle S\rangle_n = \infty$ $\mathbb{P}^\alpha$-a.s., with $\langle S\rangle_n := \sum_{k=1}^n \mathbb{E}[Z_k^2|\mathcal{F}_{k-1}]$,
iii) there exists $\varepsilon_\alpha > 0$ such that $\frac{Q_n}{\langle S\rangle_n} > \varepsilon_\alpha$ for large n, $\mathbb{P}^\alpha$-a.s.

In general, the validity of conditions i), ii), iii) might be difficult to check. In addition, the process Q_n is not predictable, thus, it is not possible to conclude that statistical functionals that depends on the mean (such as ES) satisfy the consistency property as in Definition 17 using this methodology [1]. Hence, in line with the elicitability framework, verifying the accuracy of mean-based estimates is definitely more problematic than the same problem for quantile-based forecasts. For the case of ΛVaR this is possible and all the conditions are satisfied so that the methodology can be successfully applied.

6. Conclusions

We have shown that ΛVaR, satisfies robustness and elicitability in particular classes of distributions. Robustness requires that the Λ function is continuous and does not coincide with the distribution F on any interval. Elicitability requires a bit more, that is, Λ is crossed only once by any possible F. We have also proposed an example of construction of an elicitable and robust ΛVaR given a set of normal distributions. In addition, we have shown that ΛVaR satisfies the consistency property without any conditions on the mechanism generating data, allowing a straightforward back-testing.

After the recent financial crisis, the Basel Committee (2013) has suggested that banks should abandon VaR in favour of the ES as a standard tool for risk management since ES is able to overcome two main shortcomings of VaR: lack of convexity on random variables and insensitivity with respect to tail behaviour. However, ES has also some issues. Specifically, ES is not robust, or only for small degrees when a stronger definition of robustness is required, and it is not elicitable. Recently, Acerbi and Székely (2014) showed that the elicitability of ES can be reached jointly with VaR (see also Fissler and Ziegel 2016, for an extended result). In addition, verifying the consistency property for ES is more problematic. Moreover, a recent study by Koch-Medina and Munari (2016) pointed out that not all the aspects of ES are well understood. For instance, for

[1]We thank an anonymous referee that pointed out this issue.

positions with a high probability of losses but also high expected gains in the tails, ES does not necessarily perform better than VaR from a liability holders' perspective. Other risk measures which consider the magnitude of losses beyond ES are the expectiles, recently studied by Bellini and Di Bernardino (2015).

In any case, the issue of capturing tail risk remains crucial and cannot be accomplished through VaR. The new risk measure, ΛVaR, may solve this issue since it is able to discriminate the risk among distributions with the same quantile but different tail behaviour and shares with VaR other important properties such as quasi-convexity. On the other hand, ΛVaR lacks subadditivity and the flexibility introduced by the Λ function requires additional criteria for determining its upper and lower bound. However, we think that ΛVaR may be considered as an alternative risk measure valuable for further studies.

Acknowledgement

We wish to thank the two anonymous referees for useful comments and J. Corbetta for helpful discussions on this subject.
ETH foundation is gratefully acknowledged for supporting this research.

References

Acerbi, C. and Székely, B., Backtesting Expected Shortfall. In *Risk*, 2014.

Artzner, P., Delbaen, F., Eber, J.-M. and Heath., D., Coherent measures of risk. In *Mathematical Finance*, 1999, **9**, 203–228.

Basel Committee on Banking Supervision, Fundamental review of the trading book, second consultative document. Bank for International Settlements, 2013.

Bellini, F. and Bignozzi, V., Elicitable risk measures. In *Quantitative Finance*, 2015, **15**.

Bellini, F. and Di Bernardino, E., Risk management with expectiles. In *The European Journal of Finance*, 2015, forthcoming.

Corbetta, J. and Peri, I., Backtesting Lambda Value at Risk. Working paper, 2016. Available at arxiv.org/abs/1602.07599.

Cont, R., Deguest, R. and Scandolo, G., Robustness and sensitivity analysis of risk measurement procedures. In *Quantitative Finance*, 2010, **10**, 593–606.

Davis, M.H.A., Verification of internal risk measure estimates. In *Statistics & Risk Modeling with Applications in Finance and Insurance*, 2016, **33**, 67–93.

Drapeau, S. and Kupper, M., Risk Preferences and their Robust Representation. In *Mathematics of Operations Research*, 2012, 28–62.

Embrechts, P. and Hofert, M., A note on generalized inverses. In *Mathematical Methods of Operations Research*, 2013, **77**, 423–432.

Embrechts, P. and Hofert, M., Statistics and Quantitative Risk Management for Banking and Insurance. In *Annual Review of Statistics and Its Application*, 2014, **1**, 493–514.

Embrechts, P., Puccetti, G., Rüschendorf, L., Wang, R. and Beleraj, A., An Academic Response to Basel 3.5. In *Risks*, 2014, **2**, 25–48.

Fissler, T. and Ziegel, J., Higher order elicitability and Osbands principle. In *The Annals of Statistics* , 2016, **44**, 1680–1707.

Frittelli, M., Maggis, M. and Peri, I., Risk measures on P(R) and Value at Risk with Probability/Loss function. In *Mathematical Finance*, 2014, **24**, 442–463.

Gneiting, T., Making and evaluating point forecasts. In *Journal of the American Statistical Association*, 2011, **106**, 746–762.

Hampel, F., Ronchetti, E., Rousseeuw, P. and Stahel, W., *Robust Statistics: The Approach Based on Influence Functions*, New York, 1986.

Huber, P., *Robust Statistics*, Wiley, NJ., 1981.

Koch-Medina, P. and Munari, C., Unexpected shortfalls of Expected Shortfall: Extreme default profiles and regulatory arbitrage. In *Journal of Banking & Finance*, 2016, **62**, 141–151.

Krätschmer, V., Schied, A. and Zähle, H., Qualitative and infinitesimal robustness of tail-dependent statistical functionals. In *Journal of Multivariate Analysis*, 2012, **103**, 35–47.

Krätschmer, V., Schied, A. and Zähle, H., Comparative and qualitative robustness for law invariant risk measures. In *Finance and Stochastics*, 2014, **18**, 271–295.

Lambert, N., Pennock, D. M. and Shoham, Y., Eliciting properties of probability distributions. In *Proceedings of the 9th ACM Conference on Electronic Commerce*, Chicago, Illinois, 8-12 July 2008.

Osband, K. H., Providing incentives for better cost forecasting. PhD thesis, University of California, Berkeley, 1985.

Williams, D., *Probability with Martingales*, Cambridge University Press, 1991.

Ziegel, J., Coherency and elicitability. In *Mathematical Finance*, 2014. Available at doi: 10.1111/mafi.12080.